Online VQEG Meeting Minutes

June 7 to 11, 2021

Monday 7th June 2021

Meeting schedule confirmed.

Online introductions.

Project summaries:

* 5G Key Performance Indicators (5G-KPI). Pres VQEG\_5GKPI\_2021\_115
* Audiovisual HD (AVHD)
* Computer Generated Imagery (CGI)
  + [VQEG\_CGI\_2021\_139](https://docs.google.com/presentation/d/1uA6AXpod7tnn6glV0TotCzxk-hpk4t0EEYDfCgHxgSQ/edit?usp=sharing)
* Immersive Media Group (IMG)
* Implementer's Guide for Video Quality Metrics (IGVQM)
* JEG-Hybrid
* No Reference Metrics (NORM)
* Psycho-Physiological Quality Assessment (PsyPhyQA)
* Quality Assessment for Computer Vision Applications (QACoViA)
* Quality Assessment for Health Applications (QAH)
  + [VQEG\_QAH\_2021\_143](https://drive.google.com/file/d/1BldMw4gEeUPIvGZrNNTl2WOrMP2WWmVo/view?usp=sharing)
* Statistical Analysis Methods (SAM)

Support Groups

* eLetter
* Independent Lab Group (ILG)
* Tools and Subjective Lab Setup: <https://vqeg.github.io>
* VQEG Administration and Web Support
  + Display presentation number on slides
  + Use file name from presentation list
  + Email presentations to [mpinson@ntia.doc.gov](mailto:mpinson@ntia.doc.gov), to make available on VQEG ftp site
  + Connection information for interim meetings is on the main VQEG webpage

Other Organizations

* ITU
* MPEG: [ISO - ISO/IEC JTC 1/SC 29 - Coding of audio, picture, multimedia and hypermedia information](https://www.iso.org/committee/45316.html)
* QUALINET <http://www.qualinet.eu> Meeting 14 June 2021 9-12 CEST
* AOM
* IEEE immersive video. Longer report on Friday

JEG-Hybrid Session

Dr. Yendo Hu, “A Benchmarking Standard to bring Quality, Bandwidth, and Latency into a Common Measurement Domain”

#101, VQEG\_JEG-Hybrid\_2021\_101\_prop\_qbl\_5.pptx

Suggestion by Ioannis that the work could be of interest for the 5G-KPI Working Group.

Enrico Masala/Lohic Fotio Tiotsop, “Updates on recent JEG-Hybrid activity”

#111, VQEG\_JEG-Hybrid\_2021\_111

Reported on currently ongoing activity in the JEG-Hybrid group.

Margaret: suggested to look at datasets: ITS4S2 (images) and CCRIQ.

Steve Goering : could also share individual scores for the PVSs for their work

<https://www.researchgate.net/publication/338201010_AVT-VQDB-UHD-1_A_Large_Scale_Video_Quality_Database_for_UHD-1>  . Also check the paper <https://ieeexplore.ieee.org/document/9355144> where they also do some distribution prediction

AVHD Session

**Note: this session will be recorded and used as source footage for AVHD projects**

Steve Göring, “AVrate Voyager: an open source online testing platform”

#105, VQEG\_AVHD\_2021\_105

<https://github.com/Telecommunication-Telemedia-Assessment/AVrateVoyager>

Shahid Satti, “OPTICOM Model Performance Results”

#127, VQEG\_AVHD\_2021\_127\_OPTICOM\_VQEG\_Report\_2021-06-04-R2

Some Links/Notes from MS Teams Chat (NB):

* Staelens et al. - 2010 - Assessing quality of experience of IPTV and video on demand services in real-life environments

* there is also mode 3 available, that could maybe run on an FPGA a reference implementation is available here:<https://github.com/Telecommunication-Telemedia-Assessment/bitstream_mode3_p1204_3>

* An overview of the standardized models and their corresponding performance can also be found here ([https://ieeexplore.ieee.org/document/9234526?](https://ieeexplore.ieee.org/document/9234526?source=authoralert)

Discussion, “Live QoE”

#106

Decided to add vice-chairs to AVHD. Check if Quan Hyunh-Thu is still interested to stay on as vice-chair. Rohit is suggested for Live QoE, but wants to have an offline discussion first before deciding. Tentatively Ioannis is suggested for time being.

Tuesday June 8, 2021

Margaret raised the question on how to distribute and store the meeting videos, when anonymous ftp is closed down. Margaret confirmed that backup versions on the recorded videos can be stored at NTIA. Store space will not be an issue.

It was decided that Shahid, Nabajeet and Zafar would work on a suggestion that could be used for distribution, possibly including watermarking, on Youtube, Facebook and similar platforms.

CGI Session

Saman: introduce the speakers and presentations.

#109 part-1 Joel Jung (Tencent Media Lab), Presentation on ITU-T P.BBQCG work item

Cover the general project plan and passive test paradigm

* The work item is still open to join the project.
* The dataset will be shared among the partners, but with a fair contribution the full dataset will be available for research.

Coding conditions

* 3 codecs (H.264/HEVC/AV1). Using NVENC encoder
* 5 bitrates from 300 kbps to 100 mbps
* IPPP configuration
* 540p, 720p, 1080p, 2160p
* 30fps, 60fps, 120fps
* Bit depth 8 bit

The model block diagram is presented. Different model inputs will be considered, header info, payload info, game info, network info, control info, context info. Model Output are overall gaming QoE, estimated video quality, interaction quality, and interaction quality due to packet loss and latency.

Timeline is presented as follows:

|  |  |
| --- | --- |
| **Date** | **Task** |
| End of June 21 | Finalize crowd-sourcing document |
| End of June 21 | Content captured (database 1 + 2) and shared |
| End of July 21 | Database processed |
| End of July 21 | Crowdsourcing platform available (TUIL) |
| End of August 21 | MOS available for db 1 |
| End of Nov 21 | Submission of candidate models |
| Mid Dec 21 | MOS available for db 2 |
| End of Jan 22 | Selection of reference model, start of collaborative approach |
| End of Jan 23 | Final model available – model verification, validation db available |
| End of 23 | End of project |

Crowdsourcing approach

* Qualification task
  + Reinvite those that are eligible based on this task
* Rating task
  + The main task which including the rating tool (using 7 point continuous scale) + training + short qualification + post/pre test questionnaire + trapping questions

Evaluation criteria

* Use the classic approaches such as RMSE and PLCC
* Gaming content is special that the current codec (even the recent versions) do not take into account such as sudden change in the scene.

Zafar Shahid: in gaming the frame rate might be rendered differently, depends on the power and scene complexity.

* Joel: We did not consider it so far, but very good point to consider it.

Margaret: Who are the active partners:

* Five partners are the main editors (Tencent, Ericsson, Dolby, TU Berlin, TU Ilmenau)
* Multiple other partners are involved including the academic partners

Ioannis: Is this a joint VQEG and ITU work item and available for non ITU members to contribute and work?

    Martin (ITU): No, it is currently only ITU SG12 Work Item (unlike AVHD PNATS joint project).

#109 part-2 David Lindero (Ericsson), Presentation on ITU-T P.BBQCG work item

Cover the interactive test paradigm

Describe the test cloud gaming setup (block diagram): the decoded videos will be recorded if possible (if not affect the quality)

The  progress is a bit slower compared to passive test, as we decided to not run the test following the crowd sourcing approach (as it is not the )

Scope:

network impairment:

* manually limiting the bandwidth using netem
* Introducing the random delay and jitter
* Include patterns based on the real network patterns such as LTE, 5G.

Input:

* Keyboard and mouse
* Controller and touch screen will be considered in the second round of the test

Test parameters:

* 90 second stimulus,
* 1 game per subject to not waste time to switch between the games,
* 24 conditions in one test.

Quality factors be considered:

* Overall gaming QoE
* Smoothness of video
* Input quality scale (GIPS) covering immediate feedback, responsiveness, and controllability
* Player experience
* Player performance

Margaret: would be share the result/rating once the dataset is created.

* For now it is decided to share between the group but later we may release it similar to what has been done in P.NATS and P.NAMS.

Short presentation from Saman for discussion.

Ioannis: Question about CSF. May be use that to make sure the participants are at a proper distance from the monitor?

SAM Session

Jakub Nawała, “Reproducing the "Describing Subjective Experiment Consistency by p-Value P–P Plot" Paper”

#103, VQEG\_SAM\_2021\_103.pdf

[The paper in the ACM Digital Library](https://doi.org/10.1145/3394171.3413749)

[The paper on arXiv](https://arxiv.org/abs/2009.13372)

[The related GitHub repo](https://github.com/Qub3k/subjective-exp-consistency-check).

[ACM MM Reproducibility track](https://2021.acmmm.org/reproducibility-companion-papers)

**Lohic**: Should you first apply the tool presented or the sureal tool?

**Lohic**: Try to change the decision making based on the P–P plot to the decision making based on some numerical measure. One idea is to fit a regression line to the points on the P–P plot and calculate the distance from this regression line to the black threshold line. This distance would be a numerical indication of consistency.

Zhi Li + Rafał Figlus, “ITU-T P.913 MLE algorithm in Excel”

#136, the Excel file will be provided later.

**Lukas Krasula**: Do you use any Excel-specific functions?

**Rafał**: Mostly not. However, the “power” function is Excel-specific.

[Refresher of MLE algorithm](https://docs.google.com/presentation/d/15D9Jul2T-CtYZm5MqFIbztYDpd0kQLwS4-UuHwtGHhQ/edit#slide=id.gdf3154d396_0_313)

If you are interested in the Excel file with the MLE macro please e-mail Lucjan (ljanowsk@agh.edu.pl).

Pablo Pérez, “Subjective experiments with 4 subjects (FOWR)”

#116, VQEG\_SAM\_2021\_116

The data set will be made publicly available later.

The videos are already available on the CDVL.

[The paper on arXiv](https://arxiv.org/abs/2104.02618).

Zhi (MS Teams Chat): There is a recent book "Noise" by the Nobel laureate Daniel Kahneman. In the book, he classifies the noise in human judgement as "level noise", "pattern noise" and "occasion noise". The "level noise" is similar to the "subject bias" in our context, and the "pattern noise" and "occasion noise" combined are akin to the "subject inconsistency" term. The difference between "pattern noise" and "occasion noise" is that, for one subject, if the vote is repeated (due to internal consistency of that person, e.g. personal taste), it qualifies as "pattern noise"; if it cannot be repeated (due to mood of that person at that moment), it is qualified as "occasional noise".

Wednesday June 9, 2021

IRG AVQA Session

Perspectives on the definition of visually lossless quality for mobile and large format displays

Paper Link: <http://ri.diva-portal.org/smash/get/diva2:1255804/FULLTEXT01.pdf>

NTIA Technical Report TR-21-550: Confidence Intervals for Subjective Tests and Objective Metrics That Assess Image, Video, Speech, or Audiovisual Quality

Link: <https://www.its.bldrdoc.gov/publications/details.aspx?pub=3253>

The methods in this report could be used to compare the precision of the proposed new method with the precision of a known method (like DSCQS). The comparison requires individual subject ratings from the new method and individual subject ratings from DSCQS, for the same set of stimuli.

QACOViA Session

Chaired by Mikołaj Leszczuk (AGH University of Science and Technology, Poland) and Lu Zhang (INSA Rennes, France)

Alban MARIE, “Enhancing AI Resilience to Image Coding Artifacts Through Expert Training”

#104, file VQEG\_QACoViA\_2021\_104

Questions coming from Lohic Fotio Tiotsop (Politecnico di Torino, Italy)

Carolina Whitaker and Margaret Pinson, “Datasets to train NR metrics for computer vision applications”

#102, file VQEG\_QACoViA\_2021\_102.pptx

Datasets for quality assessment for computer vision applications (QACoViA) research are available on CDVL.

Words of appreciations coming from Lucjan Janowski (AGH University of Science and Technology, Poland) and Ioannis Katsavounidis (Facebook, USA)

CGI Session

JP Tauscher, EEG-based detection of deep fake images

#133

Applied the deepfake on a colleague and asked two groups of people to join in a test, one group familiar with the person and one group does not know the person that deep fake applied. In addition, some images from unknown people are also in the dataset.

* ERP was measured for different groups of participants for different content.
* The comparison was provided between known-real vs. know-fake as well as between known-real and mix-fake.
* The classification was done to check if different groups can be classified
* Message: when unknow-real vs. unknown-fake was applied, then classification was failed while it was working for others.

Dima:I was wondering, How were the CSPs defined? - was it purely from the measured response aggregation (and partitioning) across the test subjects, or was it based on some larger datasets, or are CSPs defined more broadly based on brain-processing models & concepts of which regions of brain correspond to specific thought-functions?

* That was based on the data that is collected measured from the participants.

5GKPI Session

Pablo Pérez, “5GKPI activity report”

#115, VQEG\_5GKPI\_2021\_115

Administrative

Discussed alternating between an in-person VQEG meeting in Spring and an online VQEG meeting in the fall/winter (when travel is more difficult).

Probably, we will need to wait until 2022 for an in-person meeting. Planning must be held far in advance.

Tentative weeks for fall VQEG meeting:

* 6 to 10 December
* 13 to 17 December, Tentative agreement
* 10 to 14 January

MPEG meets 17 to 21 January

Electronic Imaging meets 18 to 21 January, and 25 to 28 January

Week 6 to 10 December has some holidays

January preferred by people on the call

Time of day of current meeting discussed.

CGI Session

Thursday June 10, 2021

Nabajeet Barman, Saman Zadtootaghaj, “Evaluation of MPEG-5 Part 2 (LCEVC) for Gaming Video Streaming Applications”

#131, VQEG\_CGI\_2021\_131

Presentation slides: [VQEG\_CGI\_2021\_131](https://docs.google.com/presentation/d/19aTU6x9bO7NTpL2TdU6aHMr_OdO4k9i-gvDxOQoqijA/edit?usp=sharing)

Uses Cloud Gaming Video Dataset (CGDVS) from <https://drive.google.com/drive/folders/1B4B27FvOSx2XyuMHSSq0aPj3zjpguc5T>

LCEVC demo: search ‘V-Nova’ in app stores

**Credentials (only work for one day)**:

User name: LCEVC\_Gaming

Pwd: Gaming2021

Presentation shows quality - bit-rate curves. Dataset will be shared with other researchers when complete. Let Saman know if you are interested.

Objective of the work:

* valuation of the new MPEG-5 Low Complexity Enhancement Video Coding (LCEVC) standard on Gaming Content
* Comparison of compression efficiency of LCEVC with existing practical implementation of video codec standards H.264/AVC and H.265/HEVC

The result is presented for BD-BR analysis for H.265/HEVC and H.264/AVC of medium and veryfast presets.

Result:

x264 vs LCEVC\_x264 (medium preset): around average 42% saving considering VMAF as ground truth.

x265 vs LCEVC\_x265 (veryfast preset) :around average 35% saving considering VMAF as ground truth.

Computation time:

* x264 (medium) --> LCEVC x264 (medium) is  1.7-2x faster
* x265 (veryfast) --> LCEVC x265 (veryfast) is 1.2-1.3x faster

Subjective test will be conducted soon.

Contact persons from V-Nova: Simone Ferrara ([simone.ferrara@v-nova.com](mailto:simone.ferrara@v-nova.com)) and Guendalina Cobianchi ([guendalina@v-nova.com](mailto:guendalina@v-nova.com)). Should you like to receive a reference implementation please contact Simone or Guendalina.

Cosmin: do you plan to check the result for lower resolution?

* We did not have time to prepare it for this presentation, but it will be considered.

Cosmin: do you plan to use other metrics such as VMAF-NEG or SSIM.

* Good idea specially using VMAF NEG, and we will consider it. Other VQA metric values including ones presented will be made available later as part of the open source dataset to be published here: <https://github.com/NabajeetBarman/LCEVC-LiveGaming>

Nabajeet Barman, “User Generated HDR Gaming Video Streaming: Dataset, Codec Comparison and Challenges”

#134, VQEG\_CGI\_2021\_134

Presentation Link: [VQEG\_CGI\_2021\_134](https://docs.google.com/presentation/d/1mYBobuRU-CdXlxC5LDxMB95kx1t7vR7S-1j-iID70yk/edit?usp=sharing)

Objective: to make HDR and UGC dataset at 4k resolution and compare the performance of different codecs.

Link for the paper: <https://ieeexplore.ieee.org/document/9422736>

Link for the GamingHDRVideoSet: <https://github.com/NabajeetBarman/GamingHDRVideoSET>.

Content: Two sequences per game are recorded out of 9 games (total 18 sequences), with BT.2020 HDR content, encoded with HEVC. The 9 games were selected because of popularity and availability at HDR mode.

Encoding setting: 4k resolution only-- H.264. H.265, VP9 and AV1 are selected as choice of codec.

Result:

* Presented based on Quality vs Bitrate curves. As a choice of quality metric, HDR-VQM and PSNR are used.
* BD-BR analysis is presented for the two metrics.
* Presented the temporal quality variation according to the frame level quality.

Lukas: Is it realistic to have the framerate of 30 fps for gaming content and we have up to 120 fps.

* It is more due to challenges in recording which we tried to record up to 60 fps which frame drops happened.

Lukas: have you seen if there is a certain sudden temporal movement that might not be covered at 30 fps?

* We did not investigate it due to the pandemic.

Cosmin: based on the result for VQM, the H.264 performs higher compared to vp9, did I interpret it correctly?

* That is correct, we have seen that in other works, and it is more due to the implementation and selection of the settings than the standard’s performance.

NORM

Thursday June 10, 2021

Werner Robitza, “Impact of Spatial and Temporal Information on Video Quality and Compressibility”

#113, VQEG\_NORM\_2021\_113

Presentation and associated QoMEX 2021 paper analyze the accuracy of SI and TI. Proposes method to estimate the compressibility of a video sequence.

Ioannis: Rejecting few source-encode pairs seems questionable.

Alex Raake (copied from the MS Teams chat): “short clarification regarding Ioannis question about Werner's talk: for the four sequences omitted for P.1204.3 (out of 564), MOS\_subj was in [1, 1.15], so anyways in an area that may be excluded in follow-up analyses as you suggested”

Saman: In gaming applications and especially when considering long sequences (e.g., 30 s long), using the maximum TI is not a good idea.

Margaret: It’s really interesting that for MOS scores the TI provides significantly more information on compressibility than the SI.

Cosmin: suggest doing motion estimation (e.g., 8x8 blocks) to find the nearest match, then compute TI. Complexity after allowing for re-use of information.

Related Google Doc: [Spatial and Temporal Information (SITI)](https://docs.google.com/document/d/1pjAJet6YMznf1pPZ_5Xp0L3UiJvCh78x2LRIHvKApro/edit).

Further questions:

Comments noted by Werner (may be lacking some remarks):

- Ioannis:

  - Restrict scores on the lower end and repeat the analysis

  - Better restrict the range of scores to something practically deployed

  - Werner: Could also weight it depending on what is streamed (but depends on a particular service)

- Lohic:

  - SI/TI make a sequence hard to compress, but the details could be masked later on

  - For long sequences it does not make sense

  - For constant motion or certain local motion it's not so easy anymore to use these features

- Shahid:

  - Why is SI not performing so well?

  - Are SI/TI working well?

  - Is compressibility universal?

- Ioannis:

  - Why is TI75 working well, but max not?

- Cosmin:

  - TI-Min would make sense for sequences that are not still at all vs. sequences that are still

  - Generate sequences with particular characteristics and see if they are robust

  - E.g. 30fps with 60fps doubling would have TI-min==0 and break the analysis

  - Slow/faster pans would have high TI but compress well

- Lucjan:

  - Increase the number of sequences to prevent statistical issues

- Joel:

  - We should think about how a codec works! Tools and coding modes (merge mode) are now much different

- General:

  - Compressing and encoding many times takes time

  - Block-limited TI, much better than TI – e.g. 8x8 blocks

Mikołaj Leszczuk, “Content Type Indicator For Technologies Supporting Video Sequence Summarisation”

#114, VQEG\_NORM\_2021\_114

Ioannis: You may try to use the databases submitted to the challenge organised by Margaret.

Margaret: The videos and images are available on the CDVL.

Ioannis Katsavounidis, Summary and Discussion of Standard Metadata

#107

The related Google Doc: [A case for embedding video quality metrics as metadata in compressed video bitstreams](https://docs.google.com/document/d/1v02cd7tFz-YozfctAy2OkKz1dGX8eHuk-1lueaxobYE/edit#heading=h.m00y9qeje3rt).

Unfortunately, the related ISO document is behind a paywall.

Next Audio Call: 7 July, 17 CET, [NORM](https://www.its.bldrdoc.gov/vqeg/projects/no-reference-metrics-norm.aspx), standard meta data, [Connection Link](https://teams.microsoft.com/l/meetup-join/19%3ameeting_ZDNjZGUzZjctZmQ2OS00NDFmLTk3OWUtNzEwNzAyMTUwNWQ4%40thread.v2/0?context=%7b%22Tid%22%3a%22d6cff1bd-67dd-4ce8-945d-d07dc775672f%22%2c%22Oid%22%3a%229c2fb3d3-e6ec-48f8-96cf-2e3ed5e67800%22%7d)

NORM, Discussion on Compression Complexity

Shahid: if SI and TI are not offering high correlation to “area under curve” compressibility of a video, then what would be “good enough” for this purpose?

Ioannis: what do we want from “complexity”? Indication of how many bits needed to achieve quality?

Shahid: use case = help balance a subjective database. But don’t want to encode a video many times to measure this, as per Werner’s convex hull.

Margaret: use case, network to estimate curve, how likely will an increase or decrease in bandwidth impact MOS.

Ioannis: want **curve**, not number. MOS vs Bit-Rate. Maybe 4 parameters or control points in a logistics function.

Nabajeet: The BD-BR functions used by MPEG does not seem to consider the saturation points and we should probably work towards a modified function which can detect such saturation when reporting the bitrate or quality savings

Ioannis: impactful problems for standardization, saturation at high end, thus should restrict range to avoid high and low end, due to diminishing returns.

Cosmin: complex problem, interested in SI/TI subset of problem, rank order only would be potentially useful (e.g., intra only mode, motion estimation). Thus, may need to imperfectly emulate motion estimation and then estimate complexity.

Ioannis: how do this efficiently? Small modifications, reasonable complexity, then this is a good improvement. Motion estimation missing from SI/TI framework. Agnostic way, doesn’t follow standard, e.g., old method that misses refinements but core mechanism.

Shahid: might allow us to rank complexity correctly, but may not compression behavior.

Ioannis: which codec? AVC? HEVC? Goal is generic coding, reduction factor from one should be accurate for others.

Ioannis may be able to make available suitable motion estimation software.

Lucjan: SI/TI also suitable for understanding diversity of content. For complexity, maybe focus on a small area. This would lower the computational complexity problem.

Cosmin: sub-set of SI/TI discussions, advancing beyond SI/TI. Also, initial goal should be rank ordering.

Silvio: Tradeoff between complexity and accuracy. Favor less complex methods, very simple,

Goals:

* Simple algorithm
* Estimates rank order / curve
* Includes motion estimation

Ioannis: propose use X.264, open source encoder, runs very fast,

Run, time (benchmark time), compare sequence compression times,

Encodes ⇒ bitrate, compressed video, estimate quality with objective metric like VMAF

If SI/TI run not too much faster, this provides interim solution

**Agreement Reached:** Werner and Cosmin nominated and accepted as NORM Vice Chairs

AVHD Session

Ioannis Katsavounidis, “Alliance for Open Media (AOM) subjective testing validation of AV1 performance”

#129, VQEG\_AVHD\_2021\_129

Presentation includes an update and summary of AV1, including available software codecs.

Test plan motivated by MPEG test plans

Seeking:

* feedback on test plan
* Availability of labs to perform testing (esp on short notice)

Lucjan: why double stimulus?

Margaret: since want videos with no scene cuts, use 4 second sequences

Margaret: interest in public safety content, may be able to assist as a test lab but would need a lot of warning.

Saman: suggest use gaming content, offer to provide

Lucjan: public safety content use case very different, mixing with entertainment content may not work well, may be scored differently

Lukas: Will SDR and HDR content be mixed in one set? What is the plan for the viewing environment? (Ioannis: thinking separate set-ups, one for SDR, one for HDR). Lukas supports separate environments and sessions for SDR and HDR.

Lukas: fewer labels better (e.g., 100-level with 5 labels)

Saman: prefer 7-level or continuous scale based on gaming experience

Mikolaj: surveillance video, should change scoring as mentioned by Lucjan, task impact, may need to use different rating approach (e.g., recognize object)

Margaret: compelling paper from NTT, concluding 5-level ACR is best.

T. Tominaga, T. Hayashi, J. Okamoto and A. Takahashi, “Performance comparisons of subjective quality assessment methods for mobile video,” 2010 Second International Workshop on Quality of Multimedia Experience (QoMEX), Trondheim, 2010, pp. 82-87, <https://ieeexplore.ieee.org/document/5517948>

Lucjan: we don’t see more precision when moving from 5-level to methods with more levels. Psychological research supports this. Probably won’t gain or loose a lot moving from 5-level to a different scale. Prefers 5-level scale, because understand subject motivation.

Nabajeet: support Lucan’s reasoning and the 5-level scale. This is generic, people used to it from other applications (e.g., rate movies).

Ioannis: ACR vs Double Stimulus?

Lucjan: for small differences, DS is not much more precise unless you encounter a new video and don’t know what to look for. Not much difference in accuracy between ACR and DSIS.

Please send more feedback to Ioannis via email.

Friday June 11, 2021

IMG Session

Status of  IEEE standardization  on QoE for immersive experiences (IEEE P3333.1.4 – Light Field, and IEEE P3333.1.3, deep learning based quality assessment)

#123

Maria Martini, Sanghoon Lee,

Marta Orduna, Methodology to Assess Quality, Presence, Empathy,  Attitude, and Attention in Social VR: International Experiences Use Case

#120

A data set with three 360-degree videos. Both videos and .

The assessment happens without interrupting the playback.

The SSDQE method does affect the evaluation of the aggregate quality of the sequence.

[Link to the Marta’s paper on arXiv](https://arxiv.org/abs/2103.02550).

The other related paper: <https://dx.doi.org/10.1109/ICCE-Berlin47944.2019.8966170>

Kamil Koniuch (edited by Jakub Nawała): Is it possible for you to share the post-hoc questionnaire?

Marta (the presenter; copied from the chat): thank you Kamil! You can find all the questions of the post-questionnaire in the paper in arxiv

Jesus (copied from the chat): Lucjan, this is one of the papers were we used the methodology Marta presented (in the references there are some more with adaptive streaming and 3D stereo video): <https://ieeexplore.ieee.org/document/6263870>

Carlos Cortés, Evaluating the impact of delay on QoE in immersive interactive environments

#121

Irene Viola, Point cloud dynamic human dataset for immersive telecommunications

#125

45 unique sequences in the dataset. The dataset is large (around 2 TB). The dataset is still being updated.

The link to the dataset will be copied here later.

Pablo Cesar, A pipeline for social virtual reality

#124

Call for collaboration. There are multiple problems to solve.

One open data set on point cloud compression is already available. One other data set is in progress.

Narciso Garcia, FVV Live: A real-time free-viewpoint video system

#137

Irene Viola, An update on ITU work item P.QXM

#126

|  |
| --- |
| Jesús Gutiérrez, IMG Work Plan - What's Next?  #135 |

Discussion

AGH (Lucjan Janowski ([ljanowsk@agh.edu.pl](mailto:ljanowsk@agh.edu.pl)), Jakub Nawała, Mikołaj Leszczuk?, Kamil Koniuch? Dawid Juszka—[juszka@agh.edu.pl](mailto:juszka@agh.edu.pl)) expressed preliminary interest in joining forces with Jesus and other people in the IMG group.

Alex Raake (from the chat): Please add TU Ilmenau in the loop for the interactive topic in IMG,  [alexander.raake@tu-ilmenau.de](mailto:alexander.raake@tu-ilmenau.de) Participation may also depend on whether we can fit it in.

Kjell Brunnström expressed preliminary interest in joining forces with the IMG.

(From the chat) [16:35] Adeyemi-Ejeye, Femi Dr (Music and Media)

I am tentatively interested - [femi.ae@surrey.ac.uk](mailto:femi.ae@surrey.ac.uk)

(From the chat) [16:39] Glenn Van Wallendael (UGent-imec)

Interest from Ghent University - imec (glenn.vanwallendael@ugent.be)

Break 10 min

QAH Session

Zohaib Amjad Khan , “Laparoscopic video quality assessment using neural networks”

#112, VQEG\_QAH\_2021\_112

Rafael Rodrigues, Meriem Outtas, “Objective Quality Assessment of Medical Image and Video: Review and Challenges”

#117, VQEG\_QAH\_2021\_117

Aditja Raj, Maria Martini, “Multivariate regression-based convolutional neural network model for fundus image quality assessment”

#132, VQEG\_QAH\_2021\_132

IMG Session

Kjell Brunnström, Legibility and Readability in Augmented Reality

#118

M. Sc thesis

[1].         Eksvärd, S. and J. Falk. (2021). *Evaluating Speech-to-Text Systems and AR-glasses: A study to develop a potential assistive device for people with hearing impairments* (978-91-89385-16-0 (ISBN)), DOI: 10.23699/yedh-qn68.

<https://www.diva-portal.org/smash/get/diva2:1535453/FULLTEXT01.pdf>

QoMEX 2021 paper

[2].         Falk, J., S. Eksvärd, B. Schenkman, B. Andrén, and K. Brunnström.  (2021). *Legibility and readability in Augmented Reality*. in *International Conference on Quality of Multimedia Experience, QoMEX 2021*. Montreal, Canada (virtual event): IEEE.

|  |
| --- |
| Abdallah El Ali, Investigating the Relationship between Momentary Emotion Self-reports and Head and Eye Movements in HMD-based 360° VR Video Watching |

#122

|  |
| --- |
| Elijs Dima, Quality of Experience in Augmented Telepresence: Effects of Viewing Positions and Depth-Aiding Augmentation |

#130

Silvia Rossi, Towards behavioural analysis of 6-DoF user when consuming immersive media

#119

Yana Nehme, “Exploring Crowdsourcing for Subjective Quality Assessment of 3D Graphics”

#140, VQEG\_IMG\_2021\_140

Thanking the host

Closing the meeting